

1213.43667X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): H. SAHARA, et al  
Serial No.:  
Filed: March 18, 2004  
Title: INFORMATION PROCESSING SYSTEM AND METHOD  
Group:

LETTER CLAIMING RIGHT OF PRIORITY

Mail Stop Patent Application  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

March 18, 2004

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on **Japanese** Patent Application No.(s) **2003-387863** filed **November 18, 2003**.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP

  
\_\_\_\_\_  
Carl I. Brundidge  
Registration No. 29,621

CIB/nac  
Attachment  
(703) 312-6600

日本国特許庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2003年11月18日  
Date of Application:

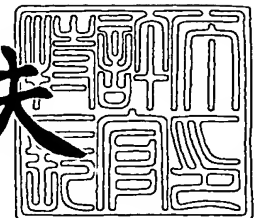
出願番号 特願2003-387863  
Application Number:  
[ST. 10/C]: [JP2003-387863]

出願人 株式会社日立製作所  
Applicant(s):

2004年 3月 5日

特許庁長官  
Commissioner,  
Japan Patent Office

今井康夫



【書類名】 特許願  
【整理番号】 GM0308020  
【提出日】 平成15年11月18日  
【あて先】 特許庁長官殿  
【国際特許分類】 G06F 3/06  
G06F 13/00  
H04L 12/00

【発明者】  
【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所  
ソフトウェア事業部内  
【氏名】 佐原 宏史

【発明者】  
【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所  
ソフトウェア事業部内  
【氏名】 鈴木 増二

【発明者】  
【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所  
ソフトウェア事業部内  
【氏名】 横内 弘

【特許出願人】  
【識別番号】 000005108  
【氏名又は名称】 株式会社日立製作所

【代理人】  
【識別番号】 100075513  
【弁理士】  
【氏名又は名称】 後藤 政喜

【選任した代理人】  
【識別番号】 100084537  
【弁理士】  
【氏名又は名称】 松田 嘉夫

【選任した代理人】  
【識別番号】 100114236  
【弁理士】  
【氏名又は名称】 藤井 正弘

【手数料の表示】  
【予納台帳番号】 019839  
【納付金額】 21,000円

【提出物件の目録】  
【物件名】 特許請求の範囲 1  
【物件名】 明細書 1  
【物件名】 図面 1  
【物件名】 要約書 1  
【包括委任状番号】 0110326

**【書類名】 特許請求の範囲****【請求項 1】**

物理デバイスに論理的に設定された論理ユニットを備える記憶装置と、  
前記記憶装置に対してデータ入出力を要求する情報処理装置と、を備え、  
前記論理ユニットに対する通信経路となる論理パスを経由してデータ入出力を要求する  
情報処理システムにおいて、  
前記データ入出力要求の対象となる前記物理デバイスの領域と連続する領域のデータを  
先読みして記憶するキャッシュメモリが設けられ、  
前記情報処理装置は、  
前記論理ユニットを分割して設けられた複数のブロックの構成を管理するパス選択情報  
管理部を備え、  
前記記憶装置に送信されるデータ入出力要求を前記論理パスに割り当てる I O 割当部と  
、  
前記データ入出力要求を所定のプロトコルにて前記 I O 割当部によって割り当てられた  
論理パスに送信する I O 処理部と、を備え、  
前記パス選択情報管理部は、少なくとも一つの前記ブロックには一つの前記論理パスを  
割り当てることを特徴とする情報処理システム。

**【請求項 2】**

前記 I O 割当部は、前記論理パスの障害を検出するパス障害検知部と、前記論理パスの  
障害の回復を検出するパス障害回復検知部とを備え、  
前記パス選択情報管理部は、前記パス障害検知部によって前記論理パスの障害が検出さ  
れ、又は、前記パス障害回復検知部によって前記論理パスの障害の回復が検出されたとき  
に、前記ブロックのサイズ及び／又は前記ブロックに割り当てるパス数を変更して、前記  
ブロックの構成を再設定することを特徴とする請求項 1 に記載の情報処理システム。

**【請求項 3】**

前記キャッシュメモリは、前記記憶装置に備わるディスク制御部に設けられることを特  
徴とする請求項 1 に記載の情報処理システム。

**【請求項 4】**

前記キャッシュメモリは、前記 I O 処理部に設けられることを特徴とする請求項 1 に記  
載の情報処理システム。

**【請求項 5】**

前記パス選択情報管理部は、  
前記情報処理装置の起動時に認識することができた前記論理パスに関する情報が記載さ  
れるパス情報管理テーブルを作成するパス情報管理部と、  
前記パス情報管理テーブルを参照して、前記論理ユニットに関する情報が記載される L  
U 情報管理テーブルを作成するブロック分割情報管理部と、  
前記 LU 情報管理テーブルを参照して、前記ブロックに関する情報が記載されるブロッ  
ク情報管理テーブルを作成するブロック情報管理部とを備え、  
前記パス選択情報管理部は、前記ブロック情報管理テーブルを参照して、前記論理ユニ  
ットを複数のブロックに分割し、前記ブロックに前記論理パスを割り当てることを特徴と  
することを特徴とする請求項 1 に記載の情報処理システム。

**【請求項 6】**

物理デバイスに論理的に設定された論理ユニットを備えた記憶装置に対して、前記論理  
ユニットに対する通信経路となる論理パスを経由してデータ入出力を要求する情報処理装  
置において、  
前記論理ユニットを分割して設けられた複数のブロックの構成を管理するパス選択情報  
管理部と、  
前記記憶装置に送信されるデータ入出力要求を前記論理パスに割り当てる I O 割当部と  
、  
前記データ入出力要求を所定のプロトコルにて前記 I O 割当部によって割り当てられた

論理パスに送信する I O 処理部と、を備え、

前記パス選択情報管理部は、少なくとも一つの前記ブロックには一つの前記論理パスを割り当てることを特徴とする情報処理装置。

【請求項 7】

前記記憶装置又は前記情報処理装置には、前記データ入出力要求の対象となる前記物理デバイスの領域と連続する領域のデータを先読みして記憶するキャッシュメモリが設けられることを特徴とする請求項 6 に記載の情報処理装置。

【請求項 8】

前記 I O 割当部は、前記論理パスの障害を検出するパス障害検知部と、前記論理パスの障害の回復を検出するパス障害回復検知部とを備え、

前記パス選択情報管理部は、前記パス障害検知部によって前記論理パスの障害が検出され、又は、前記パス障害回復検知部によって前記論理パスの障害の回復が検出されたときに、前記ブロックのサイズ及び／又は前記ブロックに割り当てるパス数を変更して、前記ブロックの構成を再設定することを特徴とする請求項 6 に記載の情報処理装置。

【請求項 9】

前記キャッシュメモリは、前記 I O 処理部に設けられることを特徴とする請求項 6 に記載の情報処理装置。

【請求項 10】

前記パス選択情報管理部は、

前記情報処理装置の起動時に認識することができた前記論理パスに関する情報が記載されるパス情報管理テーブルを作成するパス情報管理部と、

前記パス情報管理テーブルを参照して、前記論理ユニットに関する情報が記載される L U 情報管理テーブルを作成するブロック分割情報管理部と、

前記 L U 情報管理テーブルを参照して、前記ブロックに関する情報が記載されるブロック情報管理テーブルを作成するブロック情報管理部とを備え、

前記パス選択情報管理部は、前記ブロック情報管理テーブルを参照して、前記論理ユニットを複数のブロックに分割し、前記ブロックに前記論理パスを割り当てることを特徴とする請求項 6 に記載の情報処理装置。

【請求項 11】

物理デバイスに論理的に設定された論理ユニットを備えた記憶装置に対して、前記論理ユニットに対する通信経路となる論理パスを経由してデータ入出力を要求する情報処理装置において、前記論理ユニットを分割して設けられたブロックの構成を設定する情報処理装置の制御方法であって、

前記情報処理装置の起動時に、認識することができた前記論理パスに関する情報に基づいてパス情報管理テーブルを作成し、

前記パス情報管理テーブルを参照して、前記論理ユニットに関する情報が記載される L U 情報管理テーブルを作成し、

前記 L U 情報管理テーブルを参照して、前記ブロックに関する情報が記載されるブロック情報管理テーブルを作成し、

前記ブロック情報管理テーブルを参照して、前記論理ユニットを複数のブロックに分割し、

少なくとも一つの前記ブロックには一つの前記論理パスを割り当てることを特徴とする制御方法。

【請求項 12】

前記論理パスの障害が検出され、又は、前記論理パスの障害の回復が検出されたときに、前記ブロックのサイズ及び／又は前記ブロックに割り当てるパス数を変更して、前記ブロックの構成を再設定することを特徴とする請求項 11 に記載の制御方法。

【請求項 13】

物理デバイスに論理的に設定され、複数のブロックに分割された論理ユニットを備える記憶装置に対して、前記論理ユニットに対する通信経路となる論理パスを経由してデータ

入出力を要求する情報処理装置の制御方法であって、

前記データ入出力要求に基づいて、当該データ入出力要求に係るデータが存在するブロックに割り当てられている I/O 処理部を選択し、

前記選択された I/O 処理部に当該データ入出力要求を処理させる情報処理装置の制御方法。

【請求項 14】

物理デバイスに論理的に設定され、複数のブロックに分割された論理ユニットを備えた記憶装置に対して、前記論理ユニットに対する通信経路となる論理パスを経由してデータ入出力を要求する情報処理装置の制御方法であって、

前記データ入出力要求に係るデータがキャッシュメモリに記憶されているか否かを判定し、

前記データ入出力要求に係るデータが前記キャッシュメモリに記憶されていれば、前回のデータ入出力要求と同じ I/O 処理部を選択してキャッシュメモリを機能させると共に、

前記データ入出力要求に係るデータが前記キャッシュメモリに記憶されていなければ、当該データ入出力要求に係るデータが存在するブロックに割り当てられている I/O 処理部を選択し、

前記選択された I/O 処理部に当該データ入出力要求を処理させる情報処理装置の制御方法。

【請求項 15】

物理デバイスに論理的に設定された論理ユニットを備えた記憶装置に対して、前記論理ユニットに対する通信経路となる論理パスを経由してデータ入出力を要求する情報処理装置を機能させるプログラムであって、

前記情報処理装置の起動時に、認識することができた前記論理パスに関する情報に基づいて情報管理テーブルを作成する手段と、

前記パス情報管理テーブルを参照して、前記論理ユニットに関する情報が記載される L/U 情報管理テーブルを作成する手段と、

前記 L/U 情報管理テーブルを参照して、前記ブロックに関する情報が記載されるブロック情報管理テーブルを作成する手段と、

前記ブロック情報管理テーブルを参照して、前記論理ユニットを複数のブロックに分割する手段と、

少なくとも一つの前記ブロックには一つの前記論理パスを割り当てる手段として機能させることを特徴とするプログラム。

【請求項 16】

前記論理パスの障害が検出され、又は、前記論理パスの障害の回復が検出されたときに、前記ブロックのサイズ及び／又は前記ブロックに割り当てるパス数を変更して、前記ブロックの構成を再設定する手段として機能させることを特徴とする請求項 15 に記載のプログラム。

【請求項 17】

物理デバイスに論理的に設定された論理ユニットを備えた記憶装置に対して、前記論理ユニットに対する通信経路となる論理パスを経由してデータ入出力を要求する情報処理装置を機能させるプログラムであって、

前記論理ユニットを分割して設けられた複数のブロックの構成を管理する手段と、

前記記憶装置に送信されるデータ入出力要求を前記論理パスに割り当てる手段と、

前記データ入出力要求を所定のプロトコルにて前記 I/O 割当部によって割り当てられた論理パスに送信する手段と、

少なくとも一つの前記ブロックには一つの前記論理パスを割り当てる手段として機能させることを特徴とするプログラム。

**【書類名】 明細書**

**【発明の名称】** 情報処理システム、情報処理装置、情報処理装置の制御方法及びプログラム

**【技術分野】****【0001】**

本発明は、情報処理装置、情報処理装置の制御方法、プログラム及び情報処理システムに関し、特に、記憶装置からデータを読み出す際のキャッシュメモリを有効に利用する情報処理装置に関する。

**【背景技術】****【0002】**

近年、情報処理システムで取り扱われるデータ量が増大している。大容量化した記憶装置では記憶容量に見合うだけの入出力性能と信頼性を確保する必要がある。そのため記憶装置への論理パス（入出力経路）を多重化し、記憶装置へのデータ入出力要求（I/O要求）を論理パスに適宜割り当ててものが開発されている。

**【非特許文献1】** W. Curtis Preston 著，金崎裕己監訳，豊沢聡訳「SAN&NASストレージネットワーク管理」，オライリー・ジャパン，2002年10月30日，p. 66-67

**【発明の開示】****【発明が解決しようとする課題】****【0003】**

前述した従来技術では、記憶装置は、情報処理装置からのI/O要求を受信した順にI/O要求に応じた処理を実行している。

**【0004】**

しかし、記憶装置が受信するI/O要求には互いに関連するものがあり、これらのI/O要求の特性に着目し、I/O要求に応じた処理の効率化を図る仕組みは開発されていなかった。特に、負荷分散機能を機能させている場合に、I/O要求は各パスに均等に分散され、シーケンシャルI/Oであっても異なるパスに割り当てられることがある。また、記憶装置側ではシーケンシャルアクセスであることを認識することができず、記憶装置側の先読みキャッシュ機能が有効に機能せず、アクセス性能が悪くなることがあった。

**【0005】**

さらに、シーケンシャルアクセスの場合には同一パスを使用してI/O処理をする方法も提案されているが、複数のアプリケーションで同一論理ユニットにアクセスしている場合はI/O要求がシリアル化されるため、一つのアプリケーションからはシーケンシャルアクセスが要求されているとしても、論理ユニットへのアクセスは連続した領域に対するものとなるとは限らない。この場合、先読みキャッシュが有効に機能せず、アクセス性能が悪くなることがある。

**【0006】**

本発明は、シーケンシャルI/Oを同一パスで処理することによって、先読みキャッシュを有効に利用して、I/O処理を効率よく行わせることが可能な情報処理システムを提供することを目的とする。

**【課題を解決するための手段】****【0007】**

本発明は、物理デバイスに論理的に設定された論理ユニットを備える記憶装置と、前記記憶装置に対してデータ入出力を要求する情報処理装置と、を備え、前記論理ユニットに対する通信経路となる論理パスを経由してデータ入出力を要求する情報処理システムにおいて、前記データ入出力要求の対象となる前記物理デバイスの領域と連続する領域のデータを先読みして記憶するキャッシュメモリが設けられ、前記情報処理装置は、前記論理ユニットを分割して設けられた複数のブロックの構成を管理するパス選択情報管理部を備え、前記記憶装置に送信されるデータ入出力要求を、前記論理パスに割り当てるI/O割当部と、前記データ入出力要求を、所定のプロトコルにて前記I/O割当部によって割り当てら

れた論理パスに送信する I O 処理部と、を備え、前記パス選択情報管理部は、少なくとも一つの前記ブロックには一つの前記論理パスを割り当てることを特徴とする。

【発明の効果】

【0008】

本発明では、論理ユニットを分割して設けられた複数のブロックの構成を管理するパス選択情報管理部が、少なくとも一つの前記ブロックには一つの前記論理パスを割り当てるので、先読みキャッシュを有効に利用することができる。

【発明を実施するための最良の形態】

【0009】

以下、本発明の実施の形態を図面を参照して説明する。

【0010】

図 1 は、本実施の形態の情報処理システムの全体構成を示すブロック図である。

【0011】

情報処理システムは、少なくとも 1 台の情報処理装置 100 及び少なくとも 1 台の記憶装置 200 を含んで構成される。なお、本実施の形態においては、1 台の情報処理装置 100 と 1 台の記憶装置 200 とを備える情報処理システムについて説明するが、2 台以上の情報処理装置 100 及び 2 台以上の記憶装置 200 を備える情報処理システムであってもよい。

【0012】

情報処理装置 100 には、ネットワーク（例えば、SAN (Storage Area Network)）を介して記憶装置 200 が接続されている。情報処理装置 100 と記憶装置 200 との間の SAN を介する通信は、一般にファイバチャネルプロトコル (Fibre Channel Protocol) に従って行われる。すなわち、情報処理装置 100 からは、記憶装置 200 に対して、ファイバチャネルプロトコルに従ってブロック単位のデータアクセス要求が送信される。なお、情報処理装置 100 と記憶装置 200 とは、必ずしも SAN で接続される必要はなく、LAN 等のネットワークを介して接続されてもよいし、SCSI (Small Computer System Interface) インターフェースによって直接接続されてもよい。

【0013】

情報処理装置 100 と記憶装置 200 との間を接続する SAN には、複数の通信経路（論理パス）301～304 が設定されている。情報処理装置 100 は、複数の論理パス 301～304 を通じて記憶装置 200 が記憶しているデータにアクセスする。論理パス 301～304 は、情報処理装置 100 と記憶装置 200 とを接続するハードウェアによって物理的に構成されるデータ転送路である物理パスに対応づけられている。

【0014】

情報処理装置 100 は、CPU (Central Processing Unit) やメモリ等を備え、各種アプリケーションプログラムを実行する。情報処理装置 100 は、アプリケーション 110、パス管理部 120、複数の I O 処理部 131～134、複数のアダプタ 141～144 等を備えている。

【0015】

アプリケーション 110 は、情報処理装置 100 において実行される様々なアプリケーションプログラム（例えば、銀行の自動預金預け払いシステムや航空機の座席予約システム等のプログラム）である。アプリケーションプログラムは記憶装置 200 に記憶されているデータへアクセスをしながら処理を実行する。

【0016】

アダプタ 141～144 は、記憶装置 200 との間で通信を行うためのインタフェースであり、例えば、SCSI (Small Computer System Interface) アダプタ、ファイバチャネルアダプタ等の HBA (Host Bus Adapter) である。情報処理装置 100 は、アダプタ 141～144 によって、記憶装置 200 との間でデータ入出力要求 (I O 要求) 等を授受することができる。

【0017】



ＩＯ処理部１３１～１３４は、アダプタ１４１～１４４を制御するドライバやソフトウェア等（例えば、ＳＣＳＩドライバ）によって実現される。ＩＯ処理部１３１～１３４は、バス管理部１２０から通知されたＩＯ要求を順次処理する。ＩＯ処理部１３１～１３４は、ＩＯ要求を所定の通信プロトコル（ファイバチャネル、ＴＣＰ／ＩＰ等）の packets に変換し、アダプタ１４１～１４４を介して記憶装置２００に送信する。なお、以下に説明する実施の形態では、記憶装置２００側にデータバッファ２１１を設けているが、ＩＯ処理部１３１～１３４や、アダプタ１４１～１４４にデータバッファを設けてもよい。

#### 【００１８】

バス管理部１２０では、バス管理ソフトウェアが動作しており、情報処理装置１００と記憶装置２００との間の論理バス３０１～３０４を管理する。バス管理部１２０は、記憶装置２００へのＩＯ要求を送信する論理バス３０１～３０４を多重化して、帯域幅を拡大することによって、データアクセスにおけるボトルネックを解消し、データ入出力処理性能を向上させる。また、バス管理部１２０はＩＯ要求を各論理バス３０１～３０４へ適宜振り分けることによって負荷分散処理も行う。さらに、バス管理部１２０は、ある論理バス３０１～３０４に障害が発生した場合に当該論理バス３０１～３０４を切り離して、正常な論理バス３０１～３０４を通じて記憶装置２００へのデータアクセスを継続する機能も有している。これにより、論理バス障害によるサービス停止を回避しシステムの信頼性を高めている。

#### 【００１９】

なお、本実施の形態では、情報処理装置１００がコンピュータ（パーソナルコンピュータ、ワークステーション、メインフレームコンピュータ等）である場合について説明したが、情報処理装置１００が、ルータ、ハブ等のネットワーク装置であってもよい。

#### 【００２０】

記憶装置２００に記憶されているデータへの、情報処理装置１００からのアクセスは、例えば、データ書き込み要求又はデータ読み出し要求等のＩＯ要求によって行われる。このＩＯ要求は、様々なアプリケーション１１０により発行される。ＩＯ要求には、ヘッダ、記憶装置、論理ユニット（Logical Unit）、アドレス、データ長等が指定されている。ヘッダには、ＩＯ要求を送信する情報処理装置１００の識別子が記録されている。記憶装置には、ＩＯ要求の処理対象となる記憶装置の識別子が記録されている。論理ユニットには、ＩＯ要求の処理対象となる記憶装置の論理ユニットの識別子が記録されている。アドレスには、データの読み出し又は書き込み等の開始位置（先頭アドレス）が記録されている。また、データ長には、読み出し対象又は書き込み対象となるデータのサイズ（データ長）が記録されている。なお、ＩＯ要求がデータの書き込み要求である場合には、ＩＯ要求に書き込まれるデータの内容が付加される。

#### 【００２１】

アプリケーション１１０から発行されたＩＯ要求は、バス管理部１２０によって各論理バス３０１～３０４に分散され（例えば、均等に割り当てられ）、記憶装置２００に送信される。このように、複数のＩＯ要求を並列処理することによって、情報処理システムの処理性能を向上させることができる。

#### 【００２２】

記憶装置２００は、ディスク制御部２１０、論理ユニット（ＬＵ：Logical Unit）２２０等を備え、情報処理装置１００から送信されてくるＩＯ要求に応じてデータの入出力を処理する。

#### 【００２３】

論理ユニット２２０は、情報処理装置１００に提供する記憶資源（物理デバイス）を論理的に設定した記憶領域である。図１において、記憶装置２００は、一つの論理ユニットによって構成されているが、２以上の論理ユニットを備えてもよい。また、論理ユニット２２０は複数のブロック２２１～２２４に分割されており、分割されたブロック２２１～２２４に１対１に対応して論理バス３０１～３０４が設定されている。記憶資源としては、ハードディスク装置（ディスクアレイ装置）の他、フレキシブルディスク装置や半導体

記憶装置等、様々な記憶媒体を用いることができる。

#### 【0024】

ディスク制御部 210 は情報処理装置 100 から送信されてくる I/O 要求を受け、論理ユニット 220 に記憶されているデータへのアクセスを制御する。ディスク制御部 210 は、データバッファ 211 を備えている。データバッファ 211 は論理バス 301～304 に共通に設けられているが、論理バス 301～304 毎に設けてもよい。データバッファ 211 は、論理ユニット 220 に書き込まれるデータや論理ユニット 220 から読み出されたデータを一時的に記憶するキャッシュメモリとして機能する。

#### 【0025】

なお、本実施の形態においては、記憶装置 200 に先読みキャッシュ機能を備えている。この先読みキャッシュ機能とは、記憶装置 200 が受信した I/O 要求によって指定された論理ユニット 220 上の領域が、それ以前に、同一の論理バスを通じて受信した I/O 要求によって指定された論理ユニット 220 上の領域と連続している場合に、I/O 要求によって指定された論理ユニット 220 上の領域に後続する領域（シーケンシャル領域）のデータを論理ユニット 220 から先読みし、データバッファ 211 に記憶させる機能である。これにより、同一の論理バスを通じて情報処理装置 100 から論理ユニット 220 上の連続した領域を指定する I/O 要求を受信した場合に、データバッファ 211 に存在するデータを有効に利用することが可能となり、データ入出力の処理に要する時間を短縮することができる。

#### 【0026】

論理バス 301～304 を通じて情報処理装置 100 からデータ書き込み要求が送信された場合には、記憶装置 200 ではデータバッファ 211 上で読み出し（read）、変更（modify）、書き込み（write）等の処理が行われる。

#### 【0027】

例えば、論理バス 301～304 を通じて情報処理装置 100 からデータ書き込み要求が送信された場合には、記憶装置 200 はデータバッファ 211 にデータを書き込む。そして、記憶装置 200 は、データバッファ 211 へのデータの書き込み処理が終了すると、論理ユニット 220 へのデータの書き込みが終了しているか否かにかかわらず、データ書き込み要求（I/O 要求）が送信されたものと同じの論理バス 301～304 を通じて、データ書き込み要求の処理終了通知（I/O 終了通知）を、情報処理装置 100 に送信する。すなわち、情報処理装置 100 への I/O 終了通知は、論理ユニット 220 に対する実際のデータの書き込み動作とは同期することなく行われる。その後、ディスク制御部 210 は、データバッファ 211 に書き込まれたデータを論理ユニット 220 に書き込む。

#### 【0028】

また、論理バス 301～304 を通じて、情報処理装置 100 からデータ読み出し要求が送信された場合には、記憶装置 200 はデータバッファ 211 に読み出し対象のデータが存在するか否かを判定する。データバッファ 211 に読み出し対象のデータが存在する場合には、ディスク上からデータを読み出すことなく、データバッファ 211 からデータを読み出して、記憶装置 200 はデータバッファ 211 に存在する読み出し対象のデータを情報処理装置 100 に転送する。

#### 【0029】

一方、データバッファ 211 に読み出し対象のデータが存在しない場合には、ディスク制御部 210 は、データ読み出し要求の対象となる領域のデータを論理ユニット 220 から読み出して、当該読み出したデータを情報処理装置 100 に転送する。

#### 【0030】

このとき、受信した読み出し要求の対象となる論理ユニット 220 上の領域のデータを読み出すと共に、当該読み出し要求の対象となる領域と連続する領域（シーケンシャル領域）が存在する場合に、後続するシーケンシャル領域のデータ（シーケンシャルデータ）を論理ユニット 220 から読み出し、読み出したデータ（読み出し対象のデータ及びシーケンシャルデータ）をデータバッファ 211 に記憶させている。

**【0031】**

以上説明した先読みキャッシュ機能によって、論理ユニット220の同一ブロックの連続した領域に対する情報処理装置100からのI/O要求は同一の論理パスを通じてを送受信されるので、データバッファ211に存在するデータを利用することが可能となり、I/O処理の実行時間を短縮することができる。

**【0032】**

ディスク制御部210は、図1に示すように記憶装置200と一体に構成することもできるし、別体に構成することもできる。また、論理ユニットを複数のパーティションに分割して構成することもできる。

**【0033】**

図2は、図1に示した情報処理システムについてパス管理部120を詳細に表したブロック図である。

**【0034】**

パス管理部120は、I/O要求受付部121、I/O終了通知部122、I/O割り当て部123、パス選択情報管理部124、シーケンシャル判定部123、ブロック分割情報作成／通知部127及びアクセス情報表示部128等を備えている。

**【0035】**

I/O要求受付部121は、アプリケーション110によって発行されるI/O要求を受け取る。I/O終了通知部122は、I/O要求受付部121がアプリケーション110から受け取ったI/O要求に対する処理が終了した旨をアプリケーション110に通知する。

**【0036】**

シーケンシャル判定部123は、I/O要求受付部121がアプリケーション110から受け取ったI/O要求がシーケンシャルI/O要求であるかを判定する。具体的には、シーケンシャル判定部123は、受け取ったI/O要求（第二のI/O要求）に指定された記憶領域が、第二のI/O要求を受け取る前にアプリケーション110から受け取ったI/O要求（第一のI/O要求）に指定された記憶領域に後続するものであるか否かを判定する。この判定は、

- (1) 第二のI/O要求と第一のI/O要求とが指定する記憶装置が一致するか、
  - (2) 第二のI/O要求と第一のI/O要求とが指定する論理ユニットが一致するか、
  - (3) 第二のI/O要求に指定されているアドレスが、第一のI/O要求に指定されているアドレスからデータ長分進んだアドレスと一致するか、
- をデータバッファ211に記録されている全ての第一のI/O要求に対して判定する。

**【0037】**

シーケンシャル判定部123は、I/O要求が前述した(1)～(3)の全ての判定条件を満たす場合には、第二のI/O要求をシーケンシャルI/O要求であると判定する。一方、I/O要求が前述した(1)～(3)のいずれかの条件を満たさない場合には、第二のI/O要求をランダムI/O要求であると判定する。なお、このシーケンシャルI/O要求であるかの判定は、第二のI/O要求がデータ読み出し要求である場合にはデータ読み出し要求である第一のI/O要求に対して行われ、第二のI/O要求がデータ書き込み要求である場合にはデータ書き込み要求である第一のI/O要求に対して行われる。

**【0038】**

I/O割当待ちキュー126には、シーケンシャル判定部123によって判定された処理待ちのI/O要求が登録されている。

**【0039】**

I/O割当部123は、I/O割当待ちキュー126に登録されているI/O要求を割り当てるI/O処理部131～134を決定して、当該I/O要求に使用されるパス301～304を選択する。I/O処理部はパスに固有に設けられており、パスに割り当てられたI/O要求を処理し、記憶装置200に対してアクセスする。

**【0040】**

I/O割当部123によるI/O要求の割り当ては、I/Oアクセス要求に係るデータが存在

するブロック（ＩＯアクセス要求に係るデータのアドレスを含むブロック）に割り当てられるバスが選択される。また、ＩＯ割当部１２３は、負荷分散機能によって、ＩＯ要求を各バスに均等に分散させて割り当てている。例えば、ＩＯ要求を処理していないＩＯ処理部１３１～１３４、又は、ＩＯ要求の処理が終了しているＩＯ処理部１３１～１３４に割り当てる。

#### 【００４１】

また、ＩＯ割当部１２３は、バス障害検知部１２３ａ及びバス障害回復検知部１２３ｂを備えており、これらによって記憶装置２００に対するバス状態の変更を検出するバス状態確認機能を実現している。具体的には、バス障害検知部１２３ａは、ユーザがバスをoffline状態（バスに障害が発生しＩＯ要求が正常に処理できない閉塞状態）にする操作をした、又は、ＩＯ要求の処理中にバス障害を検出した場合に、バスの状態が変更されたものとして、バス情報管理部１２４ａに通知する。また、バス障害回復検知部１２３ｂは、ＩＯ処理部１３１～１３４に備わるバス状態自動回復機能によってバスの状態が回復すると、該バス障害の回復を検出して、バス情報管理部１２４ａに通知する。このバス状態自動回復機能は、バス障害検知部が有する機能であって、ユーザが稼働状態でないバスを、明示的に稼働状態にする操作を行うことなく、予め設定されたタイミング（例えば一定期間毎に稼働状態でないバスに対してＩＯ処理を行い、正常に処理できたか否かで、障害が回復しているかを確認し、自動的にバスの状態を稼働状態に変更する機能である。

#### 【００４２】

バス選択情報管理部１２４は、バス情報管理部１２４ａ、ブロック分割情報管理部１２４ｂ及びブロック情報管理部１２４ｃを備えている。バス情報管理部１２４ａは、バス情報管理テーブル２１００（図４）を作成する。ブロック分割情報管理部１２４ｂは、バス情報管理テーブル２１００を参照して、ＬＵ情報管理テーブル２２００（図５）を作成する。ブロック情報管理部１２４ｃは、ＬＵ情報管理テーブル２２００を参照して、ブロック情報管理テーブル２３００（図６）を作成する。

#### 【００４３】

ブロック分割情報作成／通知部１２７は、後述する可変分割方式において設けられ、ブロック分割情報作成／通知部１２７は、操作者の入力に基づいて、ブロック分割指定テーブル２４００を作成し、保持している。

#### 【００４４】

アクセス情報表示部１２８は、各ブロックに対するシーケンシャルアクセスＩＯ回数及びランダムアクセスＩＯ回数を出力する。例えば、利用者による設定変更に必要な情報（シーケンシャル／ランダム別のアクセス回数等）を表示する。なお、この場合、利用者は論理ユニットのブロック分割を意識して設定を変更する必要がある。

#### 【００４５】

図３は、本発明の実施の形態の各種テーブルの作成を説明する機能ブロック図である。

#### 【００４６】

バス情報管理部１２４ａは、バス管理ソフトウェアの起動時にバス情報管理テーブル２１００を作成する。バス情報管理テーブル２１００には、バス管理ソフトウェアの起動時に認識することができたバス数分のエントリが記載される。

#### 【００４７】

また、バス情報管理部１２４ａは、正常に処理されたＩＯ要求がシーケンシャルアクセスであるかランダムアクセスであるかをシーケンシャル判定部１２３が判定した結果に基づいて、シーケンシャルＩＯ回数２１０４又はランダムＩＯ回数２１０５を更新する。

#### 【００４８】

さらに、バス情報管理部１２４ａは、バス障害検知部１２３ａ又はバス障害回復検知部１２３ｂからバスの状態の変更を検出した旨の通知を受けると、バス情報管理テーブル２１００のバス状態２１０３を更新する。そして、バス情報管理部１２４ａは、バス情報管理テーブル２１００の更新を完了すると、ブロック分割情報管理部１２４ｂに対して、Ｌ

U情報管理テーブル2200の更新を要求するLU情報更新要求を送信する。また、このとき、ブロック分割情報管理部124bに対し、パスの障害が発生した又は障害が回復したパスのIDも通知する。

【0049】

なお、後述する可変分割方式においては、バス情報管理部124aは、バス情報管理テーブル2100の作成又は更新を完了すると、ブロック分割情報作成／通知部127に対して、ブロック分割指定テーブル2400の作成を要求する分割情報作成要求を送信する。

【0050】

ブロック分割情報管理部124bは、バス情報管理部124aからLU情報管理テーブル2200の作成要求を受け付けると、バス情報管理部124aが作成したバス情報管理テーブル2100を参照して、LU情報管理テーブル2200を作成する。

【0051】

また、ブロック分割情報管理部124bは、バス情報管理部124aからLU情報管理テーブル2200の更新要求を受け付けると、当該LU情報管理テーブル更新要求と共に通知された障害が発生又は回復したバスIDの情報に基づいてLU情報管理テーブル2200を更新する。

【0052】

具体的には、バス障害検知部123aからパスの障害が通知された場合は、LU情報管理テーブル2200を参照して障害が通知されたバスを抽出し、当該バスが接続されている論理ユニットを求め、LU情報管理テーブル2200の当該論理ユニットのバス数2203から1を減算する。また、該当論理ユニットのバスID欄2204から、LU情報管理テーブル更新要求によって通知された障害が発生したバスIDを削除する。

【0053】

また、バス障害回復検知部123bからバス障害の回復が通知された場合は、バス情報管理テーブル2100を参照して、障害の回復が通知されたバスが接続されている論理ユニットを求め、LU情報管理テーブル2200の当該論理ユニットのバス数2203に1を加算する。また、該当論理ユニットのバスID欄2204に、LU情報管理テーブル更新要求によって通知された障害が回復したバスIDを追加する。

【0054】

ブロック分割情報管理部124bは、LU情報管理テーブル2200の作成を完了すると、ブロック情報管理部124cに対して、ブロック情報管理テーブル2300の作成を要求する。また、ブロック分割情報管理部124bは、LU情報管理テーブル2200の更新を完了すると、ブロック情報管理部124cに対して、ブロック情報管理テーブル2300の更新を要求する。

【0055】

ブロック情報管理部124cは、ブロック分割情報管理部124bからブロック情報管理テーブル2300の作成又は更新の要求を受けると、ブロック分割情報管理部124bが作成したLU情報管理テーブル2200を参照して、ブロック情報管理テーブル2300を作成又は更新する。例えば、バス管理ソフトウェアが起動した時にブロック情報管理テーブルの作成要求が発行され、ブロック情報管理テーブル2300が作成される。

【0056】

さらに、ブロック情報管理部124cは、IO割り当て部123に対し、各ブロックにアクセスするために用いられるバスIDを通知する。

【0057】

図4には、本発明の実施の形態のバス情報管理テーブルを示す。

【0058】

バス情報管理テーブル2100は、バス管理ソフトウェアの起動時に認識することができたバスに関する情報を、常に最新の状態に更新して保持している。すなわち、バス情報管理テーブル2100は、バス管理ソフトが起動したときに新たに作成され、バスに障害

が発生したとき又はパスの障害が回復したときに情報が更新される。

【0059】

パス情報管理テーブル2100は、均等分割方式及び可変両分割方式の双方において作成される。均等分割方式は、論理ユニットを稼動状態のパス数によって均等の容量（サイズ）で複数のブロックに分割し、1ブロックに1パスを割り当てるものである。一方、可変分割方式は、ブロック分割指定テーブル2400（図7）に基づいて、論理ユニットを複数のブロックに分割するものである。

【0060】

パス情報管理テーブル2100は、パス情報管理部124aによって作成され、ブロック分割情報管理部124bがLU情報管理テーブル2200を作成（又は更新）する際に参照される。なお、パス情報管理テーブル2100は、パス管理ソフトウェアが起動している間は、常に保持されている。

【0061】

パス情報管理テーブル2100には、パスID2101、論理ユニット（LU）2102、パス状態2103、シーケンシャルIO回数2104、及び、ランダムIO回数2105の情報が含まれており、これらの情報がパスID毎に記憶されている。

【0062】

パスID2101は、パス管理部120によって、パス毎に割り当てられる識別子であり、論理パスに1対1で対応している。論理ユニット（LU）2102は、記憶装置の論理ユニット毎に割り当てられる識別子であり、一つの論理ユニットには1又は複数のパスが設定されており、一つの論理ユニットに対して1又は複数のパスIDが対応する。

【0063】

パス状態2103は、パスID2101に記載される各パスが使用可能であるのか、使用不可能であるのかを示している。すなわち、当該パスがIO要求が正常に処理できる稼動状態であれば、“online”と記載されている。一方、当該パスに障害が発生し、IO要求が正常に処理できない閉塞状態であれば、“offline”と記載されている。

【0064】

シーケンシャルIO回数2104は、シーケンシャル判定部123がIO要求をシーケンシャルアクセスかランダムアクセスかを判定した結果、当該パスで処理されたシーケンシャルアクセスIOの合計回数が記載されている。ランダムIO回数2105は、シーケンシャル判定部123がIO要求をシーケンシャルアクセスかランダムアクセスかを判定した結果、当該パスで処理されたランダムアクセスIOの合計回数が記載されている。なお、シーケンシャルIO回数2104及びランダムIO回数2105の初期値は“0”に設定される。

【0065】

図5には、本発明の実施の形態のLU情報管理テーブルを示す。

【0066】

LU情報管理テーブル2200は、各論理ユニットの状態を保持しており、ブロック分割情報管理部124bによって、パス情報管理テーブル2100を参照して作成（又は更新）される。LU情報管理テーブル2200は、パス情報管理テーブル2100が作成（又は更新）された後に、ブロック情報管理部124cがブロック情報管理テーブル2300を作成（又は更新）する際に参照される。なお、LU情報管理テーブル2200は、均等分割方式及び可変両分割方式の双方において作成される。

【0067】

LU情報管理テーブル2200には、論理ユニット（LU）2201、論理ユニットの記憶容量（LUサイズ）2202、稼動状態となっているパス数2203、及び、稼動状態になっているパスID2204の情報が含まれており、これらの情報が論理ユニット毎に記憶されている。

【0068】

論理ユニット（LU）2201は、パス情報管理テーブル2100に記載された論理ユ

ニット 2102 について、同一 LU に関する情報の重複を排除して記載したものである。論理ユニットの記憶容量 (LU サイズ) 2202 は、論理ユニット 2201 に記載されている論理ユニットの記憶容量である。

【0069】

パス数 2203 は、論理ユニット 2201 に記載された論理ユニットに接続されているパスのうち稼動状態となっているパスの総数である。このパス数 2203 は、パス情報管理テーブル 2100 のパス状態 2103 が稼動状態 (online) となっているパスの数を集計して求められる。なお、パス数 2203 に記載される稼動状態となっているパス数は、各パスの障害の発生状況に応じて減少させ、パス数を変更して、LU 情報管理テーブル 2200 に記載される。

【0070】

パス ID 2204 は、論理ユニット 2201 に記載された論理ユニットに接続されているパスのうち稼動状態となっているパスのパス ID が全て登録されている。なお、可変分割方式においては、LU 情報管理テーブル 2200 にパス ID 2204 を記載する必要はない。

【0071】

図 6 には、本発明の実施の形態のブロック情報管理テーブルを示す。

【0072】

ブロック情報管理テーブル 2300 は、適切なパスを選択するための情報を提供する。すなわち、アプリケーション 110 から送信された IO 要求から論理ユニットとその先頭アドレスの情報を取り出し、IO 割当部 123 がブロック単位に IO を振り分ける際に参照される。ブロック情報管理テーブル 2300 は、ブロック情報管理部 124c によって、LU 情報管理テーブル 2200 を参照して作成される。

【0073】

ブロック情報管理テーブル 2300 には、ブロック ID 2301、ブロックサイズ 2302、論理ユニット (LU) 2303、先頭アドレス 2304、パス ID 2305、シーケンシャル IO 回数 2306、及び、ランダム IO 回数 2307 の情報が含まれており、これらの情報がブロック ID 毎に記憶されている。

【0074】

ブロック ID 2301 は、パス管理部 120 によって、ブロック毎に割り当てられる識別子であり、分割されたブロックに 1 対 1 で対応している。

【0075】

ブロックサイズ 2302 は、ブロック ID 2301 に記載されるブロックの長さであり、ブロック情報管理部 124c が LU 情報管理テーブル 2200 を参照し、算出したサイズである。

【0076】

論理ユニット (LU) 2303 は、論理ユニット毎に割り当てられる識別子であり、パス情報管理テーブル 2100 に記載されているパス ID 2203 に対応する論理ユニット 2102 の識別子である。

【0077】

先頭アドレス 2304 は、ブロック情報管理部 124c によって求められた各ブロックの先頭アドレスである。

【0078】

パス ID 2305 は、当該ブロックへアクセスするために用いられるパス ID であり、ブロック情報管理部 124a から通知されるものであって、現在使用することができるパス ID のみが登録されている。なお、1 ブロックに 1 パスを割り当てる場合は、パス ID 2305 の欄は 1 つになるが、ブロックに複数のパスを割り当てる場合は、複数のパス ID 2305 欄が設けられる。

【0079】

シーケンシャル IO 回数 2306 は、シーケンシャル判定部 123 が IO 要求をシーケ

ンシヤルアクセスかランダムアクセスかを判定した結果、当該ブロック（当該パス）で処理されたシーケンシヤルアクセス I O の合計回数が記載されている。ランダム I O 回数 2307 は、シーケンシヤル判定部 123 が I O 要求をシーケンシヤルアクセスかランダムアクセスかを判定した結果、当該ブロック（当該パス）で処理されたランダムアクセス I O の合計回数が記載されている。なお、シーケンシヤル I O 回数 2306 及びランダム I O 回数 2307 の初期値は”0”に設定されている。

#### 【0080】

前述した均等分割方式においては、論理ユニットを稼動状態のパス数によって均等の容量（サイズ）で複数のブロックに分割するので、パス障害発生時又はパス障害回復時には、起動時と同じ手順でブロック情報管理テーブルが更新される。すなわち、パス障害が発生した場合には、ブロック数が減少するので、ブロックサイズが大きくなる。一方、パス障害が回復した場合には、ブロック数が増加するので、ブロックサイズが小さくなる（つまり、パス障害発生前の分割形態に戻る）。

#### 【0081】

図7は、本発明の別の実施の形態の各種テーブルの作成を説明する機能ブロック図である。

#### 【0082】

図7に示す実施の形態は、図3に示す実施の形態のような論理ユニット220をブロックに分割する際に、論理ユニットを各ブロックに均等に分割する均等分割方式ではなく、ブロック分割指定テーブル2400に定められたルールに基づいて、論理ユニットを複数のブロックに分割する可変分割方式である。よって、ブロック分割情報作成／通知部127には、分割される各ブロックの情報が記載されるブロック分割指定テーブル2400が保持されている。なお、均等分割方式では、ブロック分割情報作成／通知部127及びブロック分割指定テーブル2400は設けられない。

#### 【0083】

可変分割方式においては、パス管理ソフトウェアの起動時は均等分割方式によって情報処理システムの動作が開始するが、アクセス情報表示部128によって出力されるシーケンシヤルアクセス I O 回数及びランダムアクセス I O 回数に基づいて操作者が分割状態を指定し、ブロック分割情報作成／通知部127が指定内容に基づいてブロック分割指定テーブル2400を作成することによって、可変分割方式によって情報処理システムの動作が開始する。パス障害発生時又はパス障害回復時には、ブロック分割指定テーブル2400を参照し、現時点で稼動状態となっているパス数と合致する各ブロックのブロックサイズ及び各ブロックに割り当てるパス数を抽出し、ブロック情報管理テーブル2300を作成する。

#### 【0084】

なお、操作者が分割状態を指定することなく、ブロック分割情報作成／通知部127がシーケンシヤルアクセス I O 回数及びランダムアクセス I O 回数を解析することによってブロック分割指定テーブル2400を自動的に作成し、ブロック分割指定テーブル2400を最適化するようにしてもよい。

#### 【0085】

パス情報管理部124aは、図3と同様の処理によって、パス情報管理テーブル2100を作成（又は更新）して、ブロック分割情報管理部124bに対して、LU情報管理テーブル2200の作成（又は更新）を要求する。

#### 【0086】

ブロック分割情報管理部124bは、図3と同様の処理によって、パス情報管理テーブル2100を参照して、LU情報管理テーブル2200を作成又は更新して、ブロック情報管理部124cに対して、ブロック情報管理テーブル2300の作成又は更新を要求する。

#### 【0087】

ブロック情報管理部124cは、ブロック分割情報管理部124bからブロック情報管



理テーブルの作成又は更新の要求を受けると、ブロック分割情報管理部 124b が作成した LU 情報管理テーブル 2200、及び、ブロック分割情報作成／通知部 127 が作成したブロック分割指定テーブル 2400 を参照して、ブロック情報管理テーブル 2300 を作成又は更新する。

【0088】

図 8 は、本発明の実施の形態のうち、図 7 に示す可変分割方式における、各種テーブルの作成を説明する機能ブロック図である。

【0089】

ブロック分割情報作成／通知部 127 は、分割情報作成要求を受けると、ブロック分割指定テーブル 2400 を作成する。また、ブロック分割情報作成／通知部 127 は、分割情報更新要求を受けると、ブロック分割指定テーブル 2400 を更新する。

【0090】

ブロック情報管理部 124c は、ブロック分割情報作成／通知部 127 からブロック情報管理テーブルの作成又は更新の要求を受けると、ブロック分割情報管理部 124b が作成した LU 情報管理テーブル 2200、及び、ブロック分割情報作成／通知部 127 が作成したブロック分割指定テーブル 2400 を参照して、ブロック情報管理テーブル 2300 を作成又は更新する。

【0091】

図 9 には、可変分割方式（図 7、図 8）において用いられる、本発明の実施の形態のブロック分割指定テーブルを示す。

【0092】

ブロック分割指定テーブル 2400 は、アクセス情報表示部などの情報に基づいて、論理ユニットのアクセス頻度やアクセスの種別（シーケンシャルアクセスかランダムアクセスか）に応じて、自動的に（又は、ユーザによる入力に基づいて）、ブロック分割情報作成／通知部 127 によって作成される。ブロック情報管理部 124c は、ブロック分割指定テーブル 2400 が存在する場合は、ブロック分割指定テーブル 2400 を参照し、ブロック情報管理テーブル 2300 を作成する。ブロック分割指定テーブル 2400 は、使用可能な論理ユニットをどのようにブロックに分割し、各ブロックに対し現在稼働状態となっているどのパスを割り当てることが記載されている。

【0093】

ブロック分割指定テーブル 2400 には、論理ユニット（LU）2401、論理ユニットで使用可能な稼働パス数 2402、先頭アドレス 2403、ブロックに設定される割当パス数 2404 の情報が含まれており、これらの情報が論理ユニット毎に記憶されている。

【0094】

論理ユニット（LU）2401 は、パス情報管理テーブル 2100 に記載されている論理ユニット 2102 のうち、可変分割方式の対象となる論理ユニットの識別子が記載されている。

【0095】

稼働パス数 2402 は、各論理ユニットで使用可能な稼働パス数が記載されており、ブロックの分割方法を決定する際の分割情報の検索キーとなる。

【0096】

先頭アドレス 2403 は、指定されたパス数に応じてブロックを分割した際に、論理ユニット内の各ブロックに割り当てられたアドレスの先頭番地が記載されている。

【0097】

割当パス数 2404 は、各論理ユニットをブロックに分割した際に、各ブロックに割り当てる割り当てるパス数が記載されている。

【0098】

なお、IO モニタを設け、IO モニタによって IO 情報を集計して自動的に分割方法を選択してパス数を決定してもよい。例えば、ランダムアクセスが多いブロックには本発明

によるバス割り当てを適用することなく、各バスの負荷が均等になるように I/O 要求を割り当てる通常の負荷分散方法を適用することもできる。

#### 【0099】

なお、OS (operating system) で定まるパーティションに対応して、論理ユニットをブロックごとに分割するようにブロック分割指定テーブルを規定してもよい。

#### 【0100】

図 10 は、本発明の実施の形態のバス情報管理テーブル作成処理のフローチャートであり、バス管理ソフトウェアの起動時に、バス情報管理部 124a によって実行される。

#### 【0101】

まず、I/O 割当部 123 によって認識されているバス数を取得し、それぞれに固有のバス ID を割り当てる (S1001)。例えば、バスの情報を取得した順に "0" からバス ID を割り当てることができる。そして、I/O 割当部 123 から各バスが設定されている論理ユニットの ID を取得する (S1002)。このバス ID と論理ユニットの ID とは n 対 1 で対応している。

#### 【0102】

そして、取得したバス ID 及び論理ユニットの ID を、各々バス情報管理テーブル 2100 のバス ID 欄 2101 及び論理ユニット (LU) 欄 2102 に登録する (S1003)。そして、バス情報テーブルのシーケンシャル I/O 回数 2104 及びランダム I/O 回数 2105 を "0" に初期化し (S1004)、各バスの状態 2103 を "online" に初期化する (S1005)。

#### 【0103】

そして、ブロック分割情報管理部 124b に対し、LU 情報管理テーブル 2200 の作成を要求する (S1006)。

#### 【0104】

すなわち、バス情報管理部 124a は、バス管理ソフトウェアの起動時に、ホスト 100 側で認識している全てのバス、及び、そのバスを介してアクセスしている論理ユニットの情報を取得し、バス情報管理テーブル 2100 を作成し、ブロック分割情報管理部 124b に対し、LU 情報管理テーブル 2200 の作成を要求する。

#### 【0105】

図 11 は、本発明の実施の形態のバス情報管理テーブル更新処理のフローチャートであり、バス情報管理部 124a によって実行される。

#### 【0106】

バス情報管理部 124a は、バス障害検知部 123a 及びバス障害回復検知部 123b からの通知を監視している。そして、バス障害検知部 123a からバス障害発生通知を受けると (S1101)、バス情報管理テーブル 2100 の当該バスの状態 2103 を "offline" に更新する (S1102)。そして、ブロック分割情報管理部 124b に対し、LU 情報管理テーブル 2200 の更新を要求する (S1103)。なお、LU 情報管理テーブル更新要求の発行と共にバス障害が発生したバス ID も通知する。

#### 【0107】

また、バス障害回復検知部 123b からバス障害回復発生通知を受けると (S1104)、バス情報管理テーブル 2100 の当該バスの状態 2103 を "online" に更新する (S1105)。そして、ブロック分割情報管理部 124b に対し、LU 情報管理テーブル 2200 の更新を要求する (S1106)。なお、LU 情報管理テーブル更新要求の発行と共にバス障害が回復したバス ID も通知する。

#### 【0108】

図 12 は、本発明の実施の形態の LU 情報管理テーブル更新処理のフローチャートであり、ブロック分割情報管理部 124b によって実行される。

#### 【0109】

ブロック分割情報管理部 124b は、バス情報管理部 124a が発行した LU 情報管理テーブル更新要求を受信すると、バス障害の発生に基づく通知であるか否かを判定する (

S1201)。この判定は、LU情報管理テーブル更新要求に障害発生／回復の情報を含めるか、又は、パス情報管理テーブル2100を参照して、LU情報管理テーブル更新要求によって通知されたパスIDを用いて当該パスの状態を取得することによって行う。

【0110】

その結果、パス障害の発生に基づく通知であれば、LU情報管理テーブル2200のパス数を”1”減算する(S1202)。そして、LU情報管理テーブル更新要求と共に通知されたパス障害が発生したパスIDをパスID欄2204から削除する(S1203)。そして、ブロック情報管理部124cに対し、ブロック情報管理テーブル2300の更新を要求する(S1204)。

【0111】

また、ブロック分割情報管理部124bは、パス情報管理部124aが発行したLU情報管理テーブル更新要求を受信すると、パス障害の回復に基づく通知であるか否かを判定する(S1205)。その結果、パス障害の発生に基づく通知であれば、LU情報管理テーブル2200のパス数を”1”加算する(S1206)。

【0112】

そして、LU情報管理テーブル更新要求と共に通知されたパス障害が回復したパスIDをパスID欄2204に追加する(S1207)。このとき、パス情報管理テーブル2100を参照して、パス障害が回復したパスIDがどの論理ユニットに関連するのかの情報を取得する。

【0113】

そして、ブロック情報管理部124cに対し、ブロック情報管理テーブル2300の更新を要求する(S1208)。

【0114】

図13は、本発明の実施の形態のブロック情報管理部124cの動作を示すフローチャートである。

【0115】

ブロック情報管理部124cは、ブロック分割情報管理部124bからのブロック情報管理テーブル作成要求を監視している。そして、ブロック情報管理テーブル作成要求を受信すると(S1301)、LU情報管理テーブル2200を参照して、各論理ユニットに設定されている稼動状態のパス数、及び、各論理ユニットの容量(サイズ)を取得する(S1302、S1303)。その後、ブロック分割指定テーブル2400が存在するか否かを判定する(S1304)。

【0116】

ブロック分割指定テーブル2400が存在すれば、可変分割方式であると判定し、ブロック分割指定テーブル2400に記載された内容に従って分割するブロック数を決定し、ブロックIDを割り当てる(S1305)。例えば、ブロック分割指定テーブル2400に記載された先頭アドレスの順に”0”からブロックIDを割り当てることができる。

【0117】

そして、ブロック分割指定テーブル2400を参照して、各ブロックの先頭アドレスを決定する(S1306)。よって、各ブロックは当該ブロックの先頭アドレスと隣接するブロックの先頭アドレスとによって定まる容量に分割される。その後、各論理ユニットに設定されているパスを各ブロックに割り当てて、各ブロックにパスIDを割り当てる(S1307)。

【0118】

一方、ブロック分割指定テーブル2400が存在していなければ、均等分割方式であると判定し、各論理ユニットを稼動状態のパス数によって複数のブロックに分割し、ブロックIDを割り当てる(S1308)。そして、論理ユニットの容量を分割するブロック数(パス数)で除算して、各ブロックのサイズを算出する(S1309)。よって、各ブロックは均等の容量に分割される。

【0119】

そして、S1309にて算出した各ブロックのサイズに基づいて、各部ブロックの先頭アドレスを決定する(S1310)。そして、当該論理ユニットに設定されているパスを各ブロックに割り当てて、各ブロックにパスIDを割り当てる(S1311)。

#### 【0120】

すなわち、ブロック情報管理部124cは、パス管理ソフトウェアの起動時に、ブロック分割情報管理部124bからブロック情報管理テーブルの作成要求を受け付けると、LU情報管理テーブル2200(必要に応じてブロック分割指定テーブル2100)を参照し、ブロック情報管理テーブル2300を作成する。均等分割方式であって、ブロック分割指定テーブル2400が作成されていない場合は、各論理ユニットを稼動状態のパス数分で均等に分割してブロック情報管理テーブル2300を作成する。なお、均等分割方式において、ブロック情報管理テーブルを作成(又は更新)するために必要な情報は全てLU情報管理テーブルに格納されている。

#### 【0121】

以上説明した図10、図12、図13に示す処理が、パス管理ソフトウェアの起動時に順次実行され、パス情報管理テーブル2100、LU情報管理テーブル2200及びブロック情報管理テーブル2300が作成される。

#### 【0122】

図14は、本発明の実施の形態のIO処理のフローチャートであり、パス管理部120によって実行される。

#### 【0123】

IO要求受付部121がIO要求を受け付けると、シーケンシャル判定部123は、当該IO要求がシーケンシャルアクセスIOかランダムアクセスIOかを判定する(S1401、S1402)。この判定は、当該IO要求より先に行われた同種のIO要求(前のIO要求)において指定された領域と連続している領域に対するアクセスかを判定する。そして判定の終了したIO要求は、IO割り当て待ちキュー126に登録される。

#### 【0124】

そして、シーケンシャル判定部123による判定の結果、シーケンシャルアクセスIOであると判定されたとき、IO割り当て部123は、当該IO要求を前のIO要求と同一のパスで処理するように当該IO要求を割り当てるIO処理部131~134を特定し、IO処理部特定情報に設定する(S1403)。このIO処理部特定情報は、論理ユニット内のアドレスと、そのIO処理の際に使用したパスIDを記憶したテーブルであって、要求されたIO処理の開始アドレスが、テーブルに記載のアドレスに続くもの(又は、特定の範囲内のもの)であれば、シーケンシャルアクセスであると判定し、このテーブル(IO処理部特定情報)に記載のパスIDでIOを処理する。正常にIO処理が完了した場合は、テーブルに記載されているアドレスに、今回のIO処理によって処理されたデータサイズを加える。なお、使用したパスIDはテーブルに記載されているパスIDと一致するため、このパスIDを更新する必要はない。また要求されたIO処理の開始アドレスが、テーブルに記載されているアドレスと連続するもの(あるいは特定の範囲内のもの)でなければ、ランダムアクセスであると判定し、ブロック情報管理テーブル2300を参照して使用するパスを求める。

#### 【0125】

一方、シーケンシャル判定部123による判定の結果、ランダムアクセスIOであると判定されたとき、当該IO要求から論理ユニットの情報及びアクセスすべき先頭アドレスの情報を抽出する(S1404)。なお、ランダムアクセスIOかシーケンシャルアクセスIOかを判定することなく(すなわち、S1401~S1403の処理を行うことなく)、すべてのIO要求に対してS1404の処理を実行し、IO要求から論理ユニットの情報及びアクセスすべき先頭アドレスの情報を抽出してもよい。

#### 【0126】

そして、ブロック情報管理テーブル2300を参照し、論理ユニットの情報からアクセス対象となる論理ユニットを特定し、先頭アドレスからアクセス対象となるブロックを特

定する。そして、ブロック情報管理テーブル 2300 のパス ID 欄から当該ブロックに割り当てられているパスを特定する。すなわち、当該ブロックに割り当てられているパスに I/O 要求を送信するための I/O 処理部 131~134 を特定し、I/O 処理部特定情報を設定する (S1405)。

【0127】

その後、特定された I/O 処理部に I/O 要求を送り、当該特定された I/O 処理部 131~134 によって当該 I/O 要求を処理する (S1406)。そして、当該 I/O 処理が成功したか否かを判定する (S1407)。そして、当該 I/O 処理が成功すればこの処理を終了する。一方、当該 I/O 処理が成功しなければ、バス障害検出部 123a が障害が発生したパス ID を検出し、当該パス ID をバス情報管理部 124 に通知する (S1408)。

【0128】

以上説明したように、本発明の実施の形態では、論理ユニットをブロックに分割し、データ入出力要求の対象となる物理デバイスの領域と連続する領域のデータを先読みして記憶するキャッシュメモリが設けられており、同一ブロックへのシーケンシャルアクセスは同一の論理パスを使用するので、負荷分散機能が動作しているときでも、先読みキャッシュ機能を有効に利用することができる。よって、同一ブロックへのシーケンシャルアクセスは同一パスを使用するので、負荷分散機能が動作しているときでも、先読みキャッシュ機能を有効に利用することができる。

【0129】

また、シーケンシャルアクセス中に、一時的にランダムアクセスが発生しても、そのランダムアクセスが異なるブロックへのアクセスならば、異なるパスを使用する。よって、再度前回シーケンシャルアクセスしていた領域に連続した I/O 要求であれば、前回のキャッシュ内容を利用することができる。特に、複数アプリケーションが同一の論理ユニットに同時にアクセスする場合で、異なるブロックへのシーケンシャルなアクセスが交互に要求される場合であっても、キャッシュ内容を有効に利用することができる。

【0130】

さらに、前述した効果が最大限に得られるよう、ブロック分割のサイズ、ブロックに割り当てるパス数を変更することができ、論理ユニットへのアクセス性能を向上させることができる。すなわち、ブロックサイズを変更することによって、各パスへの I/O 要求を均等化し、アクセス性能の向上を図ることができる。また、ブロックに割り当てるパス数を複数にして、ブロック内において負荷分散機能を有効に動作させることができる。このブロック内の負荷分散機能は、1 ブロックへのアクセスにランダムアクセスが多い場合に有効となる。

【0131】

また、論理パスの障害が検出され又は論理パスの障害の回復が検出されたときに、論理ユニットを分割して設けられた複数のブロックの構成を再設定する (例えば、前記ブロックに割り当てられるサイズ及び/又は前記ブロックに割り当てるパス数を変更する) ので、論理パスの状態に応じて、I/O 処理を効率よく行わせるように、ブロックの構成を設定することができる。

【図面の簡単な説明】

【0132】

【図1】 本発明の実施の形態の情報処理システムの全体構成を示すブロック図である。

【図2】 本発明の実施の形態の情報処理システムのバス管理部の詳細な構成を示すブロック図である。

【図3】 本発明の実施の形態の均等分割方式における各種テーブルの作成を説明する機能ブロック図である。

【図4】 本発明の実施の形態のバス情報管理テーブルの説明図である。

【図5】 本発明の実施の形態の LU 情報管理テーブルの説明図である。

【図6】 本発明の実施の形態のブロック情報管理テーブルの説明図である。

【図 7】本発明の実施の形態の可変分割方式における各種テーブルの作成を説明する機能ブロック図である。

【図 8】本発明の実施の形態の可変分割方式における各種テーブルの作成を説明する機能ブロック図である。

【図 9】本発明の実施の形態のブロック分割指定テーブルの説明図である。

【図 1 0】本発明の実施の形態のパス情報管理テーブル作成処理のフローチャートである。

【図 1 1】本発明の実施の形態のパス情報管理テーブル更新処理のフローチャートである。

【図 1 2】本発明の実施の形態の L U 情報管理テーブル更新処理のフローチャートである。

【図 1 3】本発明の実施の形態のブロック情報管理部の動作を示すフローチャートである。

【図 1 4】本発明の実施の形態の I O 処理のフローチャートである。

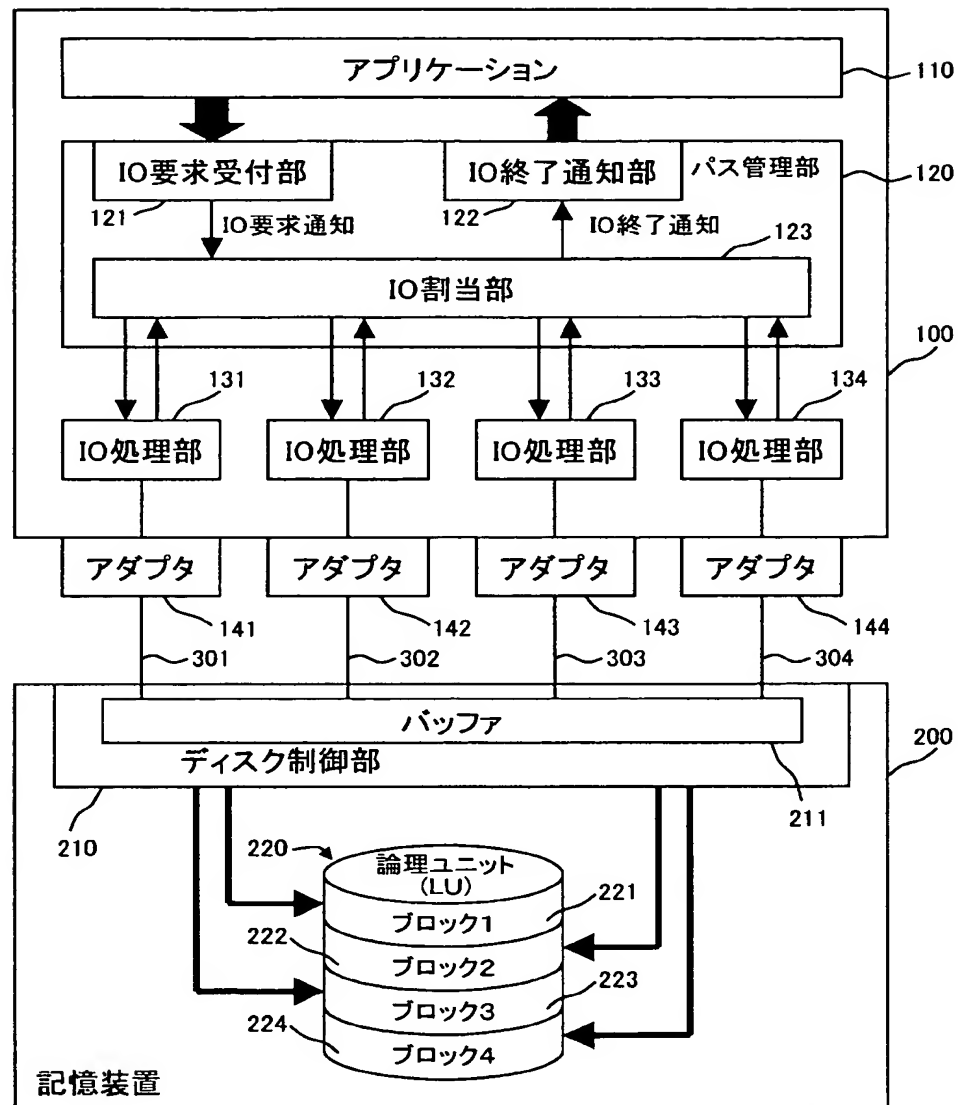
【符号の説明】

【 0 1 3 3 】

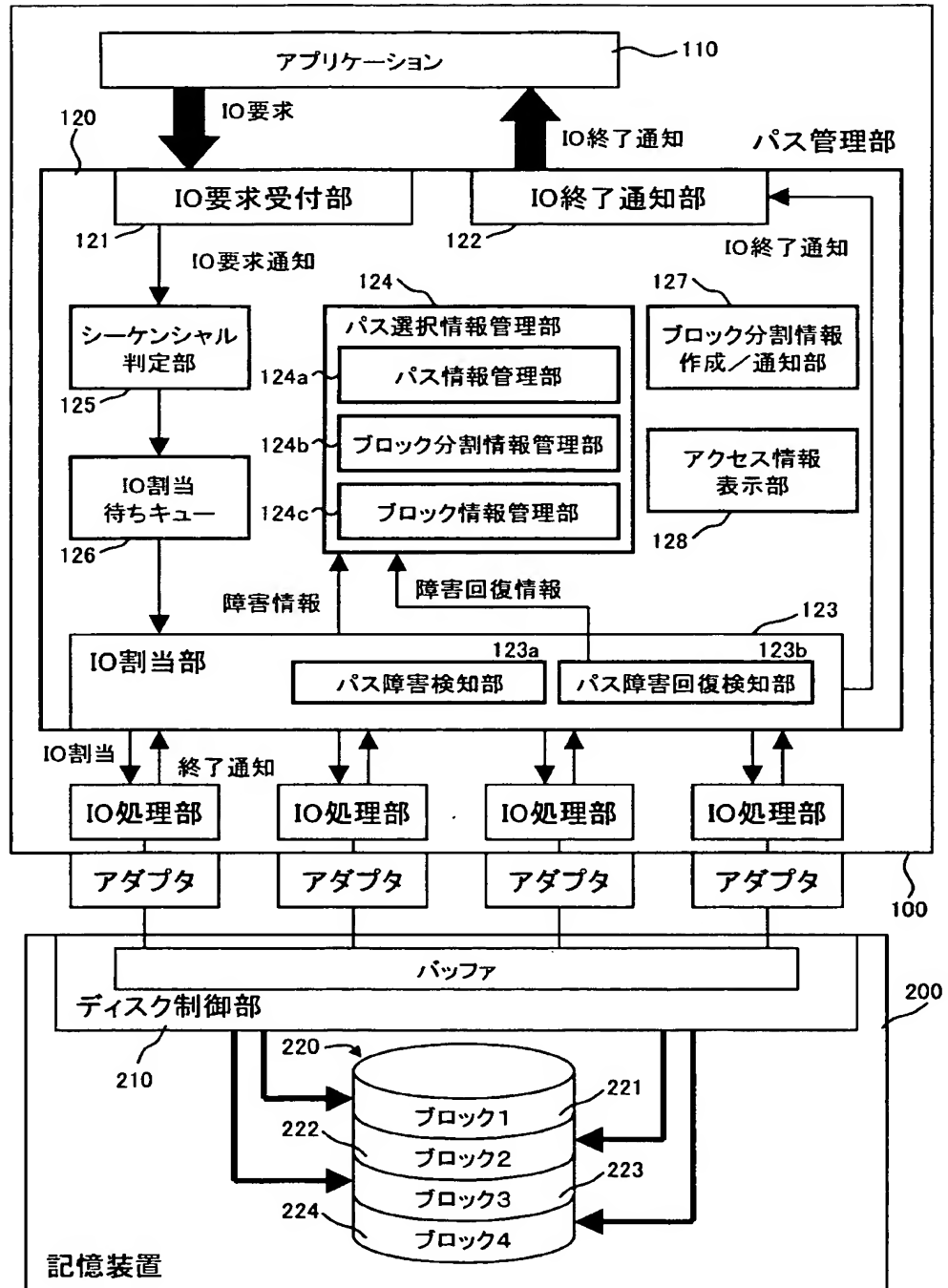
1 0 0	情報処理装置
1 1 0	アプリケーション
1 2 0	パス管理部
1 2 1	I O 要求受付部
1 2 2	I O 終了通知部
1 2 3	I O 割り当て部
1 2 4	パス選択情報管理部
1 2 4 a	パス情報管理部
1 2 4 b	ブロック分割情報管理部
1 2 4 c	ブロック情報管理部
1 2 5	シーケンシャル判定部
1 2 6	I O 割り当て待ちキュー
1 2 7	ブロック分割情報作成／通知部
1 2 8	アクセス情報表示部
1 3 1 ～ 1 3 4	I O 処理部
1 4 1 ～ 1 4 4	アダプタ
2 0 0	記憶装置
2 1 0	ディスク制御部
2 1 1	データバッファ
2 2 0	論理ユニット（論理ユニット）
3 0 1 ～ 3 0 4	論理パス

【書類名】 図面

【図 1】

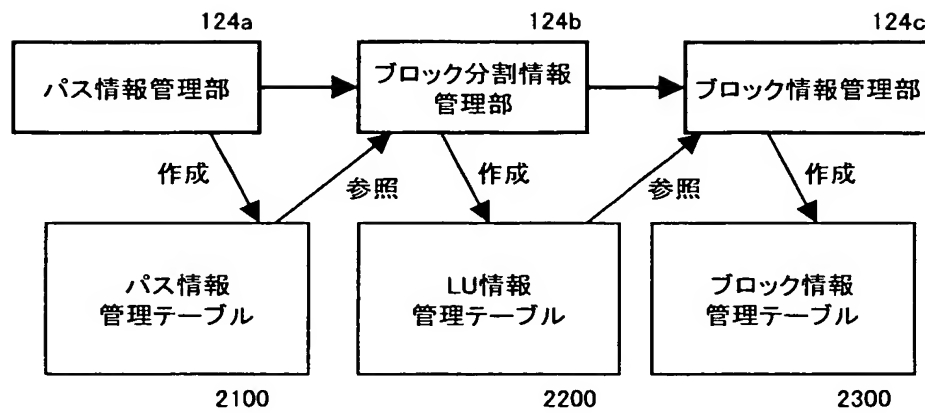


【図 2】





【図 3】



【図 4】

2101	2102	2103	2104	2105
パスID	LU	パス状態	シーケンシャルIO回数	ランダムIO回数

パス情報管理テーブル

2100

【図 5】

2201	2202	2203	2204
LU	LUサイズ	パス数	パスID

LU情報管理テーブル

2200

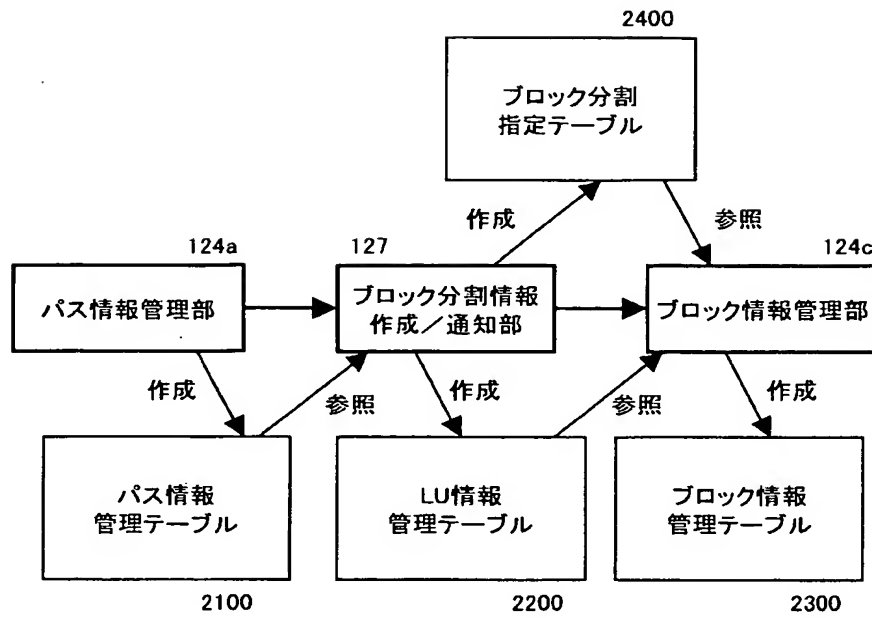
【図 6】

2301 ブロックID	2302 ブロック長	2303 LU	2304 先頭アドレス	2305 パスID	2306 シーケンシャル アクセス回数	2307 ランダム アクセス回数
1						
2						

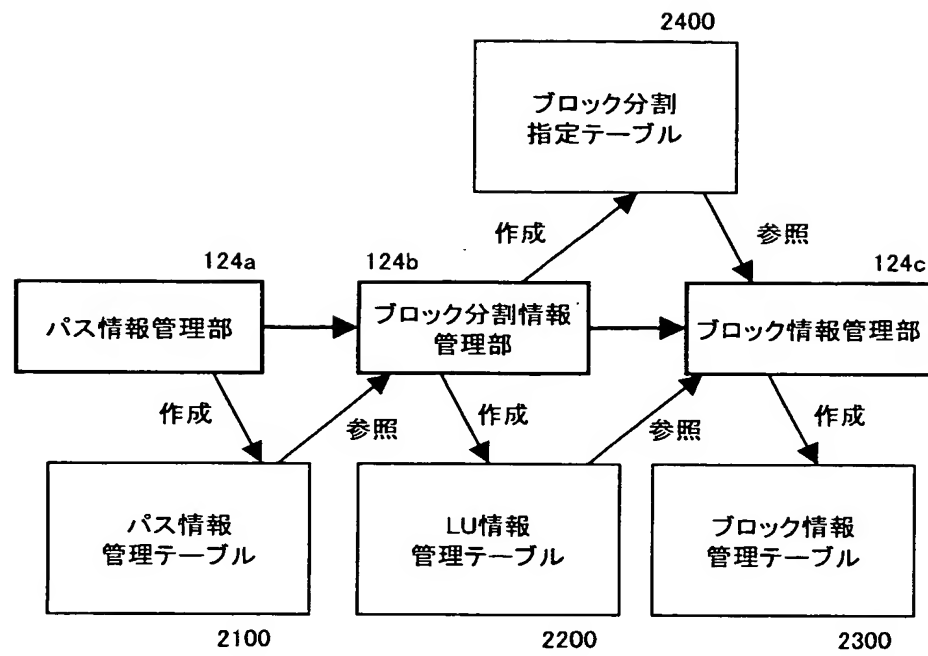
ブロック情報管理テーブル

2300

【図 7】



【図 8】



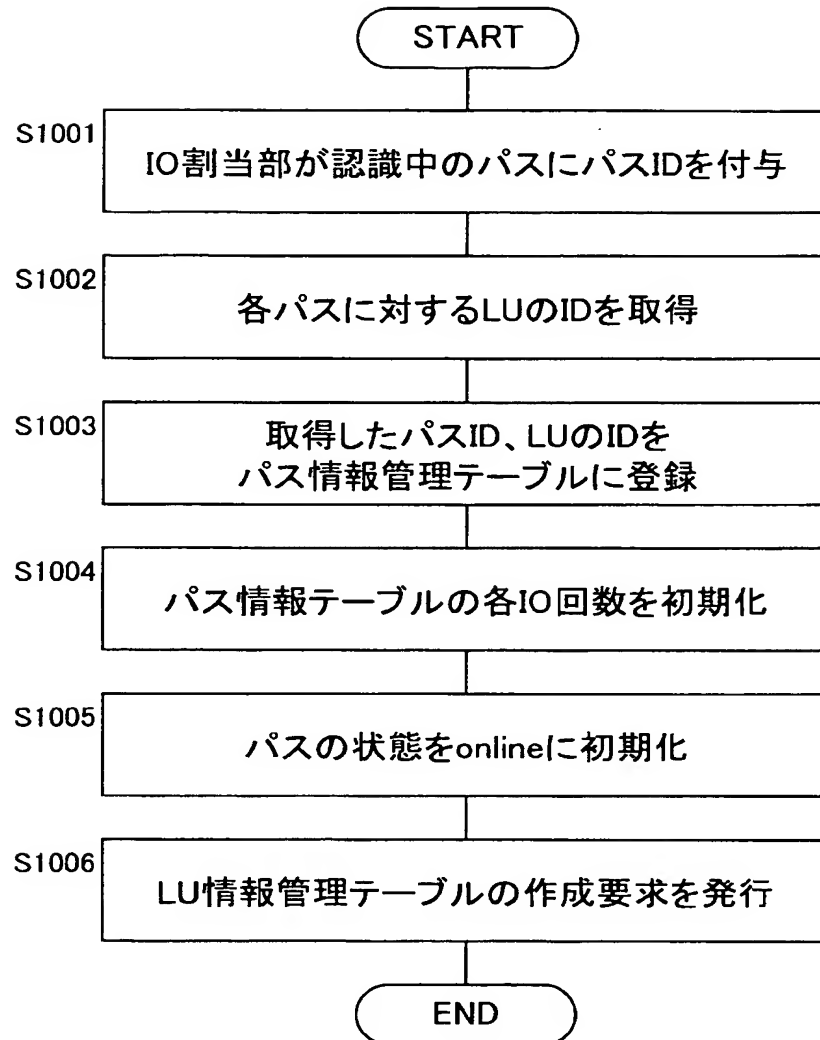
【図 9】

LU	稼働パス数	先頭アドレス	割当パス数
1	4	1,000	2
		3,000	1
		6,000	1
	3	1,000	2
		3,000	1
	2	1,000	1
		3,000	1
	1	1,000	1
2	4	1,000	4

ブロック分割指定テーブル

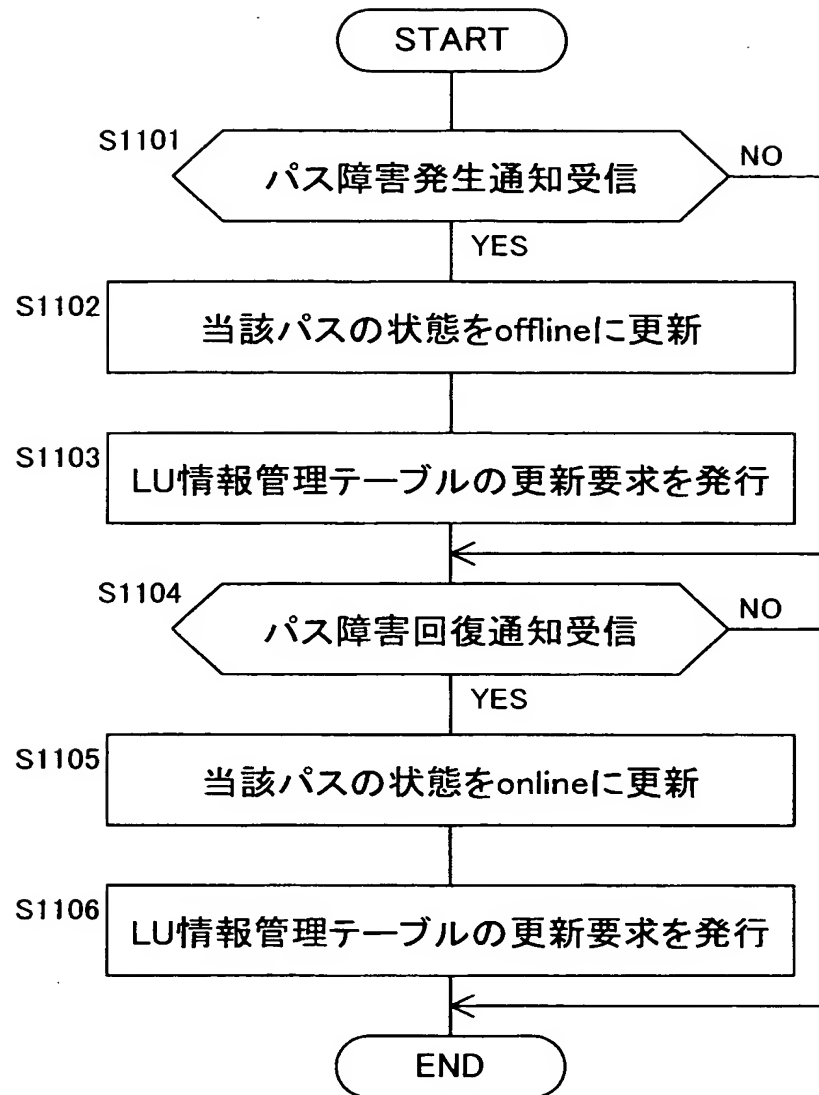
2400

【図 10】

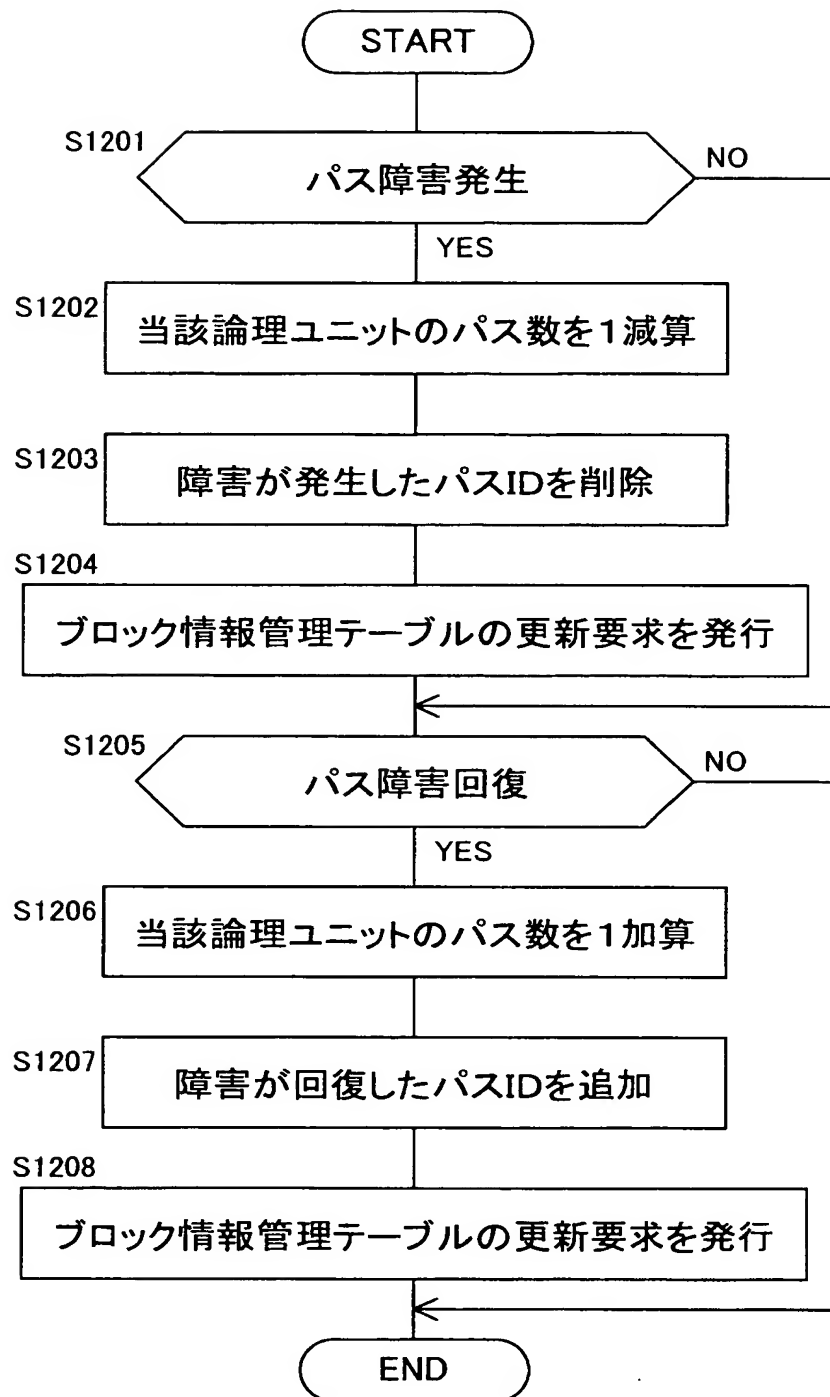




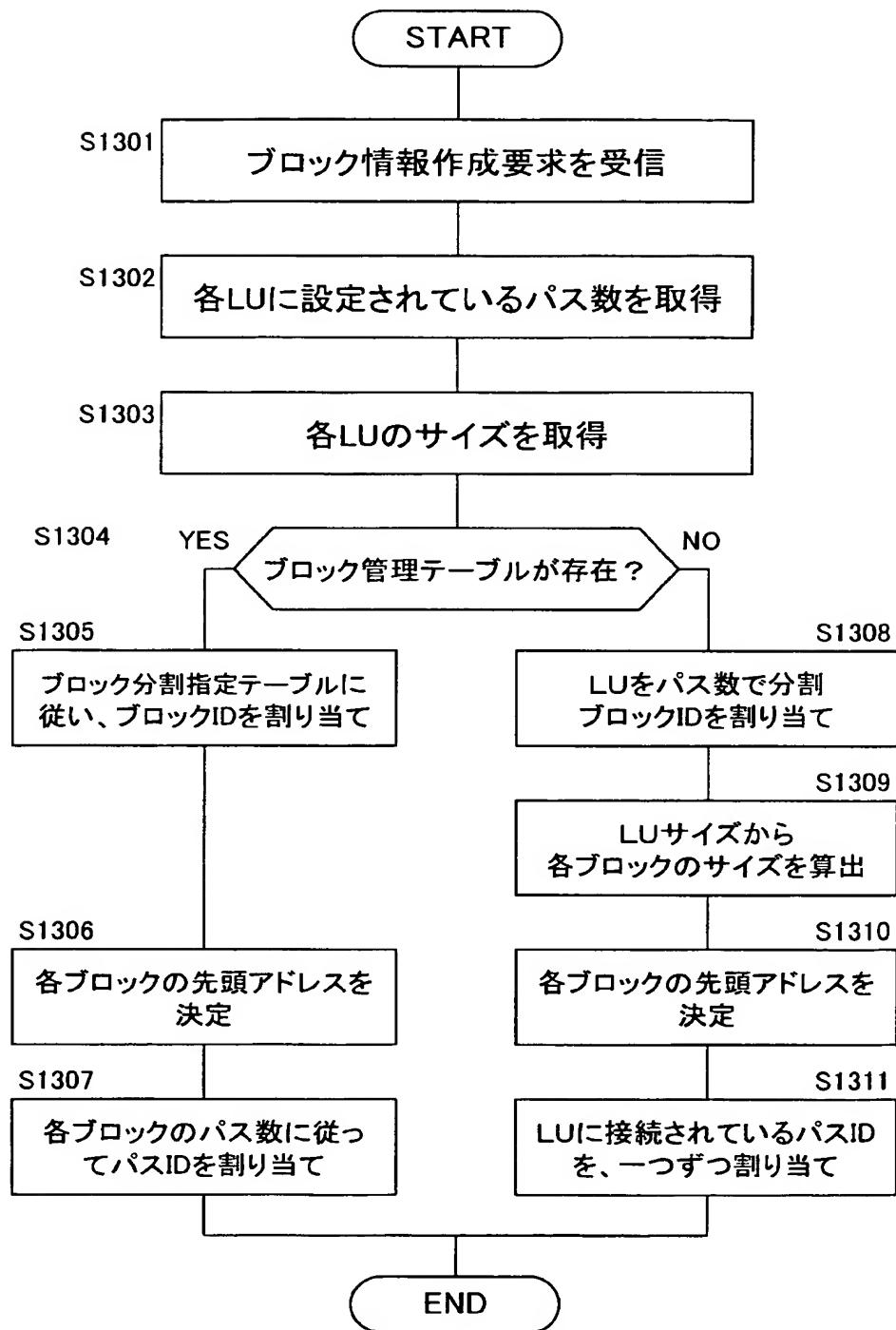
【図 11】



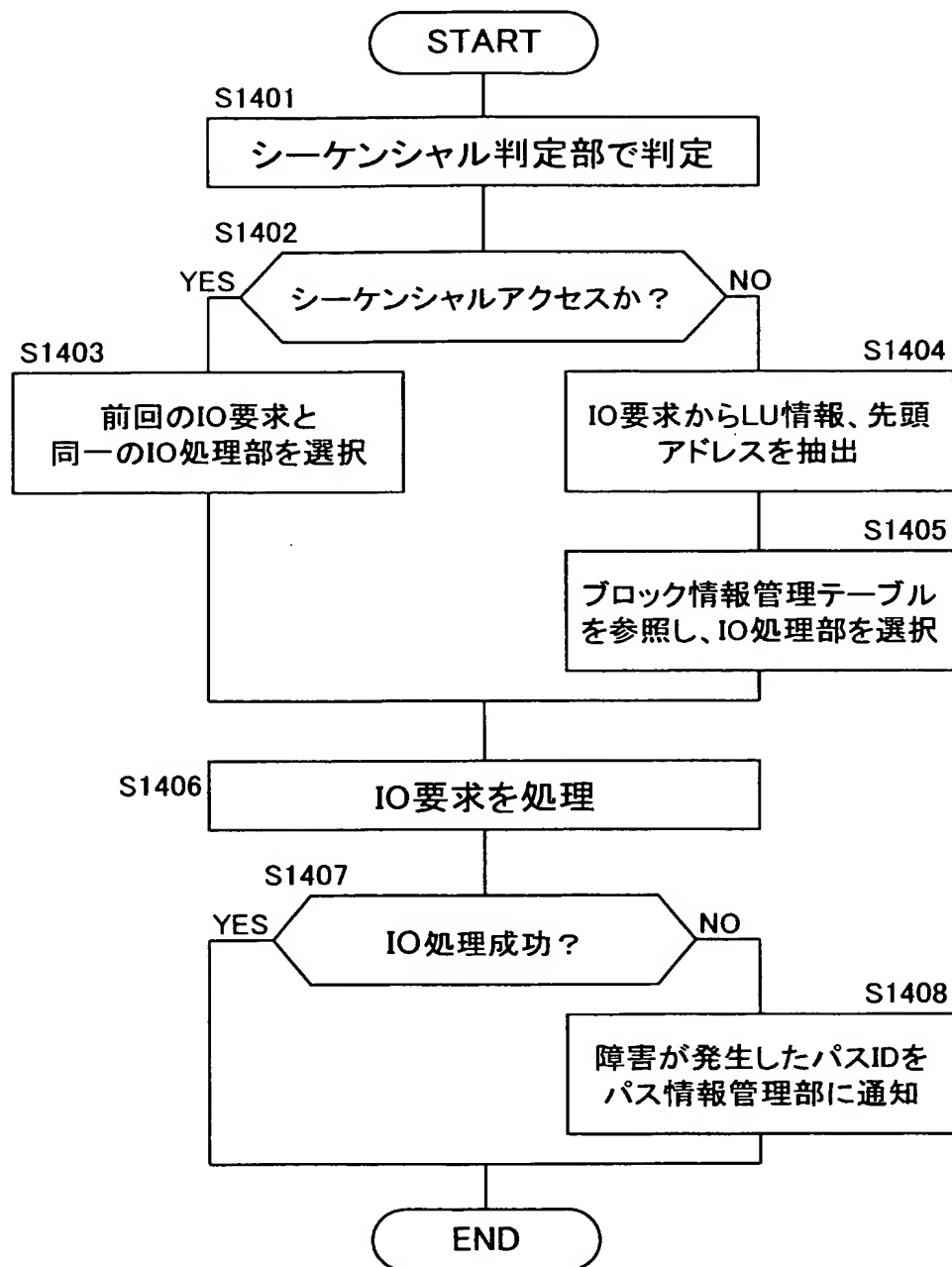
【図 12】



【図 13】



【図 14】



**【書類名】 要約書****【要約】**

**【課題】** シーケンシャル I/O を同一パスで処理することによって、先読みキャッシュを有効に利用して、効率よく I/O 要求を処理する。

**【解決手段】** 物理デバイスに論理的に設定された論理ユニット 220 を備える記憶装置 200 と、記憶装置 200 に対してデータ入出力を要求する情報処理装置 100 と、を備え、論理ユニット 220 に対する通信経路となる論理パス 301~304 を経由してデータ入出力を要求する情報処理システムにおいて、情報処理装置 100 は、論理ユニット 220 を分割して設けられた複数のブロック 221~224 の構成を管理するパス選択情報管理部 124 を備え、記憶装置 200 に送信されるデータ入出力要求を論理パスに割り当てる I/O 割当部 123 と、データ入出力要求を所定のプロトコルにて I/O 割当部 123 によって割り当てられた論理パスに送信する I/O 処理部 131~134 とを備え、パス選択情報管理部 124 は少なくとも一つのブロックには一つの論理パスを割り当てる。

**【選択図】 図 2**

特願 2 0 0 3 - 3 8 7 8 6 3

出 願 人 履 歴 情 報

識別番号 [ 0 0 0 0 5 1 0 8 ]

1. 変更年月日	1 9 9 0 年 8 月 3 1 日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台 4 丁目 6 番地
氏 名	株式会社日立製作所